International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

# Comparison between Penalty method and Lagrangian Multiplier Method

**Shelja Jhamb[1], Dr. Amanpreet Singh[2]**

*[1]Research Scholar I.K.G Punjab Technical Universiy, Jalandhar(India)*

*[2]Post Graduate Department of Mathematics,*

*SGTB Khalsa College Anandpur Sahib Punjab, (India)*

**ABSTRACT**

*In this paper we will deal with quadratic penalty function method and discuss the original multiplier method with geometric interpretation. After that existence of local minima of the Augumented Lagrangian will be discussed. Then primal function will be introduced and convergence analysis will be done for method of multipliers. In last, comparative study of Penalty method and Multiplier method will be examined.*

***Keywords: Multipliers.***

## I.THE ORIGINAL METHOD OF MULTIPLIERS

Consider the Equality Constrained Problem

(ECP) minimize $f(x)$

subject to $h(x) = 0$

where $f : R^n \to R, h : R^n \to R^m$ are given functions and $h_1,..., h_m$ are the components of $h$. For any scalar $c_k$ and $\lambda_k$ is multiplier vector. Now we will consider the Augmented Lagrangian function

$$L_{c_k}(x, \lambda_k) = f(x) + \lambda_k' h(x) + \frac{1}{2} c_k |h(x)|^2$$

In this we will assume that $x^*$ is a local minimum satisfying second order sufficient condition.

**Assumption (S):** The vector $x^*$ is a strict local minimum and a regular point of (ECP), and, $f, h \in C^2$ on some open sphere centred at $x^*$. And also $x^*$ together with its associated Lagrange multiplier vector $\lambda^*$ satisfies

$$z' \nabla^2_{xx} L_0(x^*, \lambda^*) z > 0$$

for all $z \neq 0$ with $\nabla h(x^*)' z = 0$.

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

## II.METHOD OF MULTIPLIERS

We minimize $L_{c_k}(.,\lambda_k)$ over $R^n$, given a multiplier vector $\lambda_k$ and a penalty parameter $c_k$ and then obtain a vector $x_k$. We then set

$$\lambda_{k+1} = \lambda_k + c_k h(x_k), \quad (1)$$

A penalty parameter $c_{k+1} \geq c_k$ is chosen and repeat the process.

The initial vector $\lambda_0$ is chosen arbitrarily, and the sequence $\{c_k\}$ may be either determined or pre selected on the basis of results obtained during the algorithmic process.

By using the updating formula (7), The multipliers $\lambda_k$ are determined.

## III.GEOMETRIC INTERPRETATION

Geometric interpretation of method of multipliers motivates the subsequent convergence analysis. Define the primal functional $p$ of (ECP) by

$$p(u) = \min_{h(x)=u} f(x)$$

where the minimization is local in an open sphere within which $x^*$ is the unique local minimum of (ECP). Also $p(0) = f(x^*)$ and $\nabla p(0) = -\lambda^*$. The minimization of $L_c(.,\lambda)$ can be break down into two stages, first minimizing over all $x$ such that $h(x) = u$ with $u$ fixed and then minimizing over all $u$ so that

$$\min_x L_c(x,\lambda) = \min_u \min_{h(x)=u} \left\{ f(x) + \lambda' h(x) + \frac{1}{2} c |h(x)|^2 \right\}$$
$$= \min_u \left\{ p(u) + \lambda' u + \frac{1}{2} c |u|^2 \right\}$$

where the minimization is local in the neighbourhood of $u = 0$. The minimization can be understood by the figure given below. The minimum point is found at the point $u(\lambda, c)$ and the gradient of $p(u) + \lambda' u + \frac{1}{2} c \mid u \mid^2$ will be zero at minimum point or equivalently

$$\nabla \left\{ p(u) + \frac{1}{2} c \mid u \mid^2 \right\} \Big|_{u=u(\lambda,c)} = -\lambda$$

So we get the minimum point $u(\lambda, c)$ as in figure.

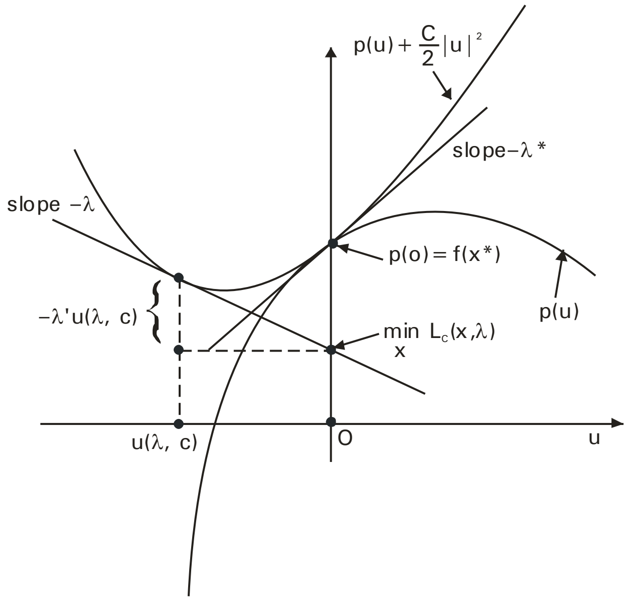


Fig. 1 Geometric interpretation of minimization of the Augumented Lagrangian

We also know that

$$\min_x L_c(x,\lambda) - \lambda' u(\lambda,c) = p[u(\lambda,c)] + \frac{1}{2}c|u(\lambda,c)|^2$$

Also the tangent of the graph $p(u) + \frac{1}{2}c|u|^2$ at $u(\lambda,c)$ intersects y-axis at $\min_x L_c(x,\lambda)$. If we take $c$ sufficiently large then $p(u) + \lambda' u + \frac{1}{2}c|u|^2$ will become convex in the neighbourhood of the origin. When $\lambda$ will close to $\lambda^*$ and $c$ will take large values then $\min_x L_c(x,\lambda)$ will be close to $p(0) = f(x^*)$.

Now we will discuss geometric interpretation of the multiplier iteration (1). Here also the vector $u_k = h(x_k)$ will minimize $p(u) + \lambda' u + \frac{1}{2}c|u|^2$ if $x_k$ minimizes $L_{c_k}(.,\lambda_k)$. Hence,

$$\nabla\left\{p(u) + \frac{1}{2}c_k|u|^2\right\}\Big|_{u=u_k} = -\lambda_k$$

And $$\nabla p(u_k) = -(\lambda_k + c_k u_k) = -[\lambda_k + c_k h(x_k)]$$

So $$\lambda_{k+1} = \lambda_k + c_k h(x_k) = -\nabla p(u_k)$$

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

as shown in figure given below

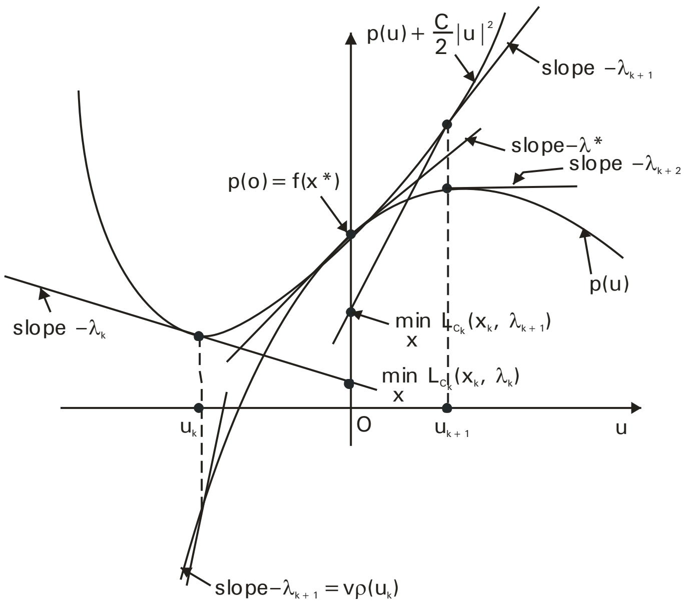


Fig. 2 Geometric interpretation of first-order multiplier iteration

the multiplier $\lambda_{k+1}$ will be closer to $\lambda^*$ than $\lambda_k$ when $c_k$ will take large values and $\lambda_k$ will be close to $\lambda^*$. Convergence to $\lambda^*$ will be obtained in one iteration when $p(u)$ will be linear and convergence will be fast enough if $\nabla^2 p(0)=0$. In this $c_k \to \infty$ is not necessary for convergence but $c_k$ must be increased at some threshold level.

## IV.EXISTENCE OF LOCAL MINIMA OF THE AUGMENTED LAGRANGIAN

In penalty method we investigate whether local minima of the augmented Lagrangian exist, and if so how far their distance from local minima of original problem is affected by the values of the penalty parameter $c$ and the multiplier $\lambda$. Here we focus on local minimum $x^*$ satisfying Assumption together with Lagrangian multiplier $\lambda^*$. For any scalar $c$ ,we have

$$L_c(x^*,\lambda^*) = f(x^*) + \lambda^* h(x^*) + \frac{1}{2}c\left|h(x^*)\right|^2$$

$$\nabla_x L_c\left(x^*,\lambda^*\right)=\nabla f\left(x^*\right)+\nabla h\left(x^*\right)\lambda^* + c\,h\left(x^*\right)\nabla h\left(x^*\right)$$

$$\nabla_x L_c\left(x^*,\lambda^*\right)=\nabla f\left(x^*\right)+\nabla h\left(x^*\right)\left[\lambda^* + ch\left(x^*\right)\right]=\nabla_x L_0\left(x^*,\lambda^*\right)=0 \qquad (2)$$

and $$\nabla^2_{xx} L_c\left(x^*,\lambda^*\right)=\nabla^2_{xx} L_0\left(x^*,\lambda^*\right)+c\nabla h\left(x^*\right)\nabla h\left(x^*\right)'$$

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

also there exists a scalar $\bar{c}$ such that

$$\nabla_{xx}^{2} L_c\left(x^*, \lambda^*\right) > 0 \quad \forall c \geq \bar{c} \tag{3}$$

hence $x^*$ is a strict local minimum of $L_c\left(., \lambda^*\right)$ for all $c \geq \bar{c}$ using (2) and (3). From above, we conclude that a local minimum of $L_c\left(., \lambda^*\right)$ close to $x^*$ exists for every $c \geq \bar{c}$, if $\lambda$ is close enough to $\lambda^*$. This will also be true when $c$ is sufficiently large even if $\lambda$ is far from $\lambda^*$. The following result makes this idea accurate.

**Proposition 1:** Let Assumption holds and let $\bar{c}$ be a positive scalar such that

$$\nabla_{xx}^{2} L_c\left(x^*, \lambda^*\right) > 0 \tag{4}$$

Then positive scalars $\delta, \varepsilon$, and $M$ exists such that:

(a) For all $(\lambda, c)$ in the set $D \subset R^{m+1}$ defined by

$$D = \left\{(\lambda, c) \,\middle|\, \left|\lambda - \lambda^*\right| < \delta c, \bar{c} \leq c\right\}, \tag{5}$$

the problem

$$\text{minimize } L_c(x, \lambda) \tag{6}$$

$$\text{subject to } x \in S\left(x^*; \varepsilon\right)$$

has a unique solution $x(\lambda, c)$. The function $x(.,.)$ is continuously differentiable in the interior of $D$, and, for all $(\lambda, c) \in D$,

$$\left|\left(x(\lambda, c) - x^*\right)\right| \leq M \left|\lambda - \lambda^*\right| / c \tag{7}$$

(b) For all $(\lambda, c) \in D$,

$$\left|\left(\tilde{\lambda}(\lambda, c) - \lambda^*\right)\right| \leq M \left|\lambda - \lambda^*\right| / c \tag{8}$$

where

$$\tilde{\lambda}(\lambda, c) = \lambda + c h[x(\lambda, c)] \tag{9}$$

(c) For all $(\lambda, c) \in D$, the matrix $\nabla_{xx}^{2} L_c[x(\lambda, c), \lambda]$ is positive definite and the matrix $\nabla h[x(\lambda, c)]$ has rank $m$.

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

Hence it can be seen that for any $\lambda$, there exists a $c_\lambda$ such that $(\lambda, c)$ belongs to $D$ for every $c \geq c_k$. The estimate $\delta c$ grows linearly with $c$ on the allowable distance of $\lambda$ from $\lambda^*$. In some cases the allowable distance may grow at a higher rate than linear. And also it is possible that for every $\lambda$ and $c > 0$, a unique global minimum of $L_c(.,\lambda)$ exists.

Thus the above proposition provide convergence and rate of convergence result for the multiplier iteration

$$\lambda_{k+1} = \lambda_k + c_k h(x_k)$$

If the generated sequence $\{\lambda_k\}$ is bounded, the penalty parameter $c_k$ is sufficiently large after a certain index, and the minimization of $L_{c_k}(.,\lambda_k)$ provides the local minimum $x_k = x(\lambda_k, c_k)$ closest to $x^*$ after that index, then we obtain $x_k \to x^*, \lambda_k \to \lambda^*$.

Now we will try to obtain a fast convergence and rate of convergence result. For this we have to introduce the primal function.

## V.THE PRIMAL FUNCTION

Take the system of equations in $(x, \lambda, u)$

$$\nabla f(x) + \nabla h(x)\lambda = 0, \qquad h(x) - u = 0$$

The above system has the solution $(x^*, \lambda^*, 0)$. Also by Implicit Function Theorem, there exists a $\delta > 0$ and functions $x(u) \in C^1$ and $\lambda(u) \in C^1$, such that $x(0) = x^*$ and $\lambda(0) = \lambda^*$, and for $|u| < \delta$,

$$\nabla f[x(u)] + \nabla h[x(u)]\lambda(u) = 0, \qquad h[x(u)] - u = 0 \tag{10}$$

Again for some and for $|u| < \delta$, we have $|x(u) - x^*| < \varepsilon$ and $|\lambda(u) - \lambda^*| < \varepsilon$. The function $p : S(0;\delta) \to R$ is given by

$$p(u) = f[x(u)] \quad \forall u \in S(0;\delta)$$

is called the primal functional corresponding to $x^*$. In view of assumptions, $x(u)$ will become actually a local minimum of the problem of minimizing $f(x)$ subject to $h(x) = u$ when we take $\delta$ and $\varepsilon$ sufficiently small.

An equivalent definition of $p$ is given by

$$p(u) = f[x(u)] = \min\{f(x) | h(x) = u, x \in S \in (x^*; \varepsilon)\} \tag{11}$$

Also $$\nabla p(u) = -\lambda(u) \quad \forall u \in S(0;\delta) \tag{12}$$

By differentiating (10), we get

$$\nabla_u x(u) \nabla_{xx}^2 L_0[x(u), \lambda(u)] + \nabla_u \lambda(u) \nabla h[x(u)]' = 0, \tag{13}$$

$$\nabla_u x(u) \nabla h[x(u)] = I \tag{14}$$

for any $c \in R$, we have, multiply (14) with $c\nabla h[x(u)]'$ both sides, then we get

$$c\nabla_u x(u) \nabla h[x(u)] \nabla h[x(u)]' = c\nabla h[x(u)]' \tag{15}$$

By adding (13) and (15), we get

$$\nabla_u x(u)\{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\} + [\nabla_u \lambda(u) - cI]\nabla h[x(u)]' = 0$$

By multiplying the above equation with $\{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\}^{-1}$ we get that for every $c$ the inverse below exists

$$\nabla_u x(u) + [\nabla_u \lambda(u) - cI]\nabla h[x(u)]' \times \{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\}^{-1} = 0$$

Multiplying both sides with $\nabla h[x(u)]$, we get

$$\nabla_u x(u)\nabla h[x(u)] + [\nabla_u \lambda(u) - cI]\nabla h[x(u)]' \times \{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\}^{-1}\nabla h[x(u)] = 0$$

Now using (12) and (14), we get

$$I + [-\nabla^2 p(u) - cI]\nabla h[x(u)]' \times \{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\}^{-1}\nabla h[x(u)] = 0$$

$$I = -[-\nabla^2 p(u) - cI]\nabla h[x(u)]' \times \{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\}^{-1}\nabla h[x(u)]$$

$$\frac{I}{\nabla h[x(u)]' \times \{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\}^{-1}\nabla h[x(u)]} = [\nabla^2 p(u) + cI]$$

$$cI + \nabla^2 p(u) = \{\nabla h[x(u)]'\{\nabla_{xx}^2 L_0[x(u), \lambda(u)] + c\nabla h[x(u)]\nabla h[x(u)]'\}^{-1}\nabla h[x(u)]\}^{-1} \tag{16}$$

Above equation holds for all $u$ with $|u| < \delta$ and for all $c$ for which the inverse above exists.

For $u=0$, we get

$$\nabla^2 p(0)=\left\{\nabla h\left(x^*\right)'\left[\nabla_{xx}^2 L_c\left(x^*,\lambda^*\right)\right]^{-1}\nabla h\left(x^*\right)\right\}^{-1}-cI \tag{17}$$

for any $c$ for which inverse of $\nabla_{xx}^2 L_c\left(x^*,\lambda^*\right)$ exists. If $\left[\nabla_{xx}^2 L_0\left(x^*,\lambda^*\right)\right]^{-1}$ exists then

$$\nabla^2 p(0)=\left\{\nabla h\left(x^*\right)'\left[\nabla_{xx}^2 L_0\left(x^*,\lambda^*\right)\right]^{-1}\nabla h\left(x^*\right)\right\}^{-1} \tag{18}$$

Now the following result will show that the point after which sudden changes are expected, can be characterized in terms of the eigen values of the matrix $\nabla^2 p(0)$.

**Proposition 2:** Suppose that the assumptions hold and for any scalar , we have

$$\nabla_{xx}^2 L_c\left(x^*,\lambda^*\right)>0 \Leftrightarrow c>\max\{-e_1,\ldots,-e_m\} \Leftrightarrow \nabla^2 p(0)+cI>0 \tag{19}$$

where $e_1,\ldots,e_m$ are the eigen values of $\nabla^2 p(0)$.

Now we shall show that the rate of convergence of the method of multipliers can also be characterized in terms of eigen values.

## VI.CONVERGENCE ANALYSIS

In this we will find the convergence and fast rate of convergence result for the method of multipliers.

**Proposition 3:** Suppose Assumption holds and let $\bar{c}$ and $\delta$ be as described above. For all $(\lambda,c)$ in the set $D$ defined by (11), there holds

$$\tilde{\lambda}(\lambda,c)-\lambda^*=\int_0^1 N_c\left[\lambda^*+\varsigma\left(\lambda-\lambda^*\right)\right]\left(\lambda-\lambda^*\right)d\varsigma \tag{20}$$

where for all $(\lambda,c)\in D$, the $m\times m$ matrix $N_c$ is defined by

$$N_c(\lambda)=I-c\nabla h[x(\lambda,c)]'\left\{\nabla_{xx}^2 L_c[x(\lambda,c),\lambda]\right\}^{-1}\nabla h[x(\lambda,c)] \tag{21}$$

where $I$ is the identity matrix.

**Theorem 4:** Suppose assumption holds, let $\bar{c}$ and $\delta$ be as described above. $e_1,\ldots,e_m$ are the eigen values of $\nabla^2 p(0)$ given by (17) or (18). Assume also that

$$\bar{c} > \max\{-2e_1,\dots,-2e_m\} \tag{22}$$

(or $\nabla^2 p(0) + \frac{1}{2}\bar{c}I > 0$). Then a scalar $\delta_1$ with $0 < \delta_1 \le \delta$ exists such that if $\{c_k\}$ and $\lambda_0$ satisfy

$$\left|\lambda_0 - \lambda^*\right| / c_0 < \delta_1, \qquad \bar{c} \le c_k \le c_{k+1} \qquad \forall k = 0,1,\dots, \tag{23}$$

Then the sequence $\{\lambda_k\}$ generatedby $\lambda_{k+1} = \lambda_k + c_k h[x(\lambda_k, c_k)]$ (24)

is well defined as $(\lambda_k, c_k)$ belongs to the set $D$ of (5) for all $k$ so $x(\lambda_k, c_k)$ is well defined.

and also we have $\lambda_k \to \lambda^*$ and $x(\lambda_k, c_k) \to x^*$. Also if $\limsup_{k\to\infty} c_k \to c^* < \infty$ and $\lambda_k \ne \lambda^*$ for all $k$,

there holds $$\limsup_{k\to\infty} \frac{\left|\lambda_{k+1} - \lambda^*\right|}{\left|\lambda_k - \lambda^*\right|} \le \max_{i=1,\dots,m} \left|\frac{e_i}{e_i + c^*}\right|, \tag{25}$$

if $c_k \to \infty$ and $\lambda_k \ne \lambda^*$ for all $k$, there holds

$$\lim_{k\to\infty} \frac{\left|\lambda_{k+1} - \lambda^*\right|}{\left|\lambda_k - \lambda^*\right|} = 0. \tag{26}$$

**Proof**: In this we will consider the matrix $N_c$ of (21), we get

$$N_c(\lambda^*) = I - c\nabla h(x^*)' \left[\nabla^2_{xx} L_c(x^*, \lambda^*)\right]^{-1} \nabla h(x^*)$$

As we have

$$\nabla^2 p(0) = \left\{\nabla h(x^*)' \left[\nabla^2_{xx} L_c(x^*, \lambda^*)\right]^{-1} \nabla h(x^*)\right\}^{-1} - cI$$

Then we get $$\nabla^2 p(0) + cI = \left\{\nabla h(x^*)' \left[\nabla^2_{xx} L_c(x^*, \lambda^*)\right]^{-1} \nabla h(x^*)\right\}^{-1}$$

and $$\left[\nabla^2 p(0) + cI\right]^{-1} = \left\{\nabla h(x^*)' \left[\nabla^2_{xx} L_c(x^*, \lambda^*)\right]^{-1} \nabla h(x^*)\right\}$$

so by using above, we get

$$N_c(\lambda^*) = I - c\left[\nabla^2 p(0) + cI\right]^{-1}$$

If we take $\mu_1(c), \dots, \mu_m(c)$ as eigen values of $N_c(\lambda^*)$, then we get

$$\mu_i(c) = 1 - \frac{c}{e_i + c} = \frac{e_i}{e_i + c}$$

As the eigen values of $\nabla^2 p(0)$ are $e_1,\ldots,e_m$ and $\nabla^2 p(0)+cI$ are $e_i + c$.

$\forall(\lambda,c)\in D$, from (21) we have

$$N_c(\lambda)=I-c\nabla h[x(\lambda,c)]'\{\nabla^2_{xx}L_c[x(\lambda,c),\lambda]\}^{-1}\nabla h[x(\lambda,c)]$$

And also we have

$$\nabla^2_{xx}L_c(x^*,\lambda^*)=\nabla^2_{xx}L_0(x^*,\lambda^*)+c\nabla h(x^*)\nabla h(x^*)'$$

Then $\nabla^2_{xx}L_c[x(\lambda,c),\lambda(\lambda,c)]=\nabla^2_{xx}L_0[x(\lambda,c),\lambda(\lambda,c)]+c\nabla h[x(\lambda,c)]\nabla h[x(\lambda,c)]'$

This implies

$$\nabla^2_{xx}L_c[x(\lambda,c),\lambda(\lambda,c)]=c\{c^{-1}\nabla^2_{xx}L_0[x(\lambda,c),\lambda(\lambda,c)]+\nabla h[x(\lambda,c)]\nabla h[\lambda,c]'\}$$

And finally we get

$$N_c(\lambda)=I-\nabla h[x(\lambda,c)]'\{c^{-1}\nabla^2_{xx}L_0[x(\lambda,c),\tilde{\lambda}(\lambda,c)]+\nabla h[x(\lambda,c)]\nabla h[x(\lambda,c)]'\}^{-1}\nabla h[x(\lambda,c)]$$

Hence by using Proposition 1 and the above result, we can easily see that for given any $\varepsilon_1>0$, $\delta_1\in(0,\delta]$ exists, such that $\forall(\lambda,c)\in D$ i.e. with $|\lambda-\lambda^*|/c<\delta_1$, $\bar{c}\le c$, we get

$$|N_c(\lambda)|\le|N_c(\lambda^*)|+\varepsilon_1=\max_{i=1,\ldots,m}|\mu_i(c)|+\varepsilon_1=\max_{i=1,\ldots,m}\left|\frac{e_i}{e_i+c}\right|+\varepsilon_1$$

And by (20), we get for all these pairs $(\lambda,c)$

$$|\tilde{\lambda}(\lambda,c)-\lambda^*|\le\left(\max_{i=1,\ldots,m}\left|\frac{e_i}{e_i+c}\right|+\varepsilon_1\right)|\lambda-\lambda^*| \qquad (27)$$

Hence from (22) and (23), we get that $\max_{i=1,\ldots,m}|e_i/(e_i+c)|<1$ hence if we will select $\varepsilon_1$ sufficiently small then we get for some $\rho\in(0,1)$ and all $(\lambda,c)$ with $|\lambda-\lambda^*|/c<\delta_1$ and $\bar{c}\le c$

$$|\tilde{\lambda}(\lambda,c)-\lambda^*|\le\rho|\lambda-\lambda^*|$$

So we get $\lambda_k\to\lambda^*$ and $x(\lambda_k,c_k)\to x^*$ by using (7) and (23). Also the convergence rate can also be calculated by using (27).

## VII.COMPARISON OF PENALTY METHOD AND MULTIPLIER METHOD

- In Penalty Method, the multiplier $\lambda_k\equiv$ constant, and it is necessary to increase the penalty parameter $c_k$ to infinity. But in method of multiplier, it is not necessary to increase $c_k$ to infinity in order to obtain convergence.

- Rate of convergence of method of multiplier is better than the penalty method.
- In method of multipliers, the rate of convergence is linear or super linear while in penalty method, the rate of convergence depends on the rate at which the penalty parameter is increased.

But it is important to know how one should select the initial multiplier $\lambda_0$ and the penalty parameter sequence. We should have to choose $\lambda_0$ as close as possible to $\lambda^*$. The following points may help us in selecting penalty parameter sequence:-

- The parameter $c_k$ should become larger than that point after which sudden changes are likely to occur.
- The initial parameter $c_0$ should not be too large.
- The parameter $c_k$ should not be increased too fast.
- The parameter $c_k$ should not be increased too slowly.

These conditions may be contradictory for some extent. Also for non convex problems, it is difficult to know about that point, after which changes are likely to occur for penalty parameter. Usually $c_0$ is chosen and subsequent values of $c_k$ are monotonically increased by the equation $c_{k+1} = \beta c_k$, where $\beta$ is a scalar with $\beta > 1$. The choices for $\beta$ can be taken from $\beta \in [4,10]$. By this, the threshold level i.e. that point after which sudden changes are likely to occur, for multiplier convergence will be exceeded.

We can increase $c_k$ by multiplication with a factor $\beta > 1$ only if the constraint violation, measured by $|h[x(\lambda_k, c_k)]|$ is not decreased by a factor $\gamma < 1$ over the minimization, i.e.

$$c_{k+1} = \begin{cases} \beta c_k & if \mid h[x(\lambda_k, c_k)] \mid > \gamma \mid h[x(\lambda_{k-1}, c_{k-1})] \mid, \\ c_k & if \mid h[x(\lambda_k, c_k)] \mid \leq \gamma \mid h[x(\lambda_{k-1}, c_{k-1})] \mid \end{cases}$$

$\beta = 10$ and $\gamma = 0.25$ are mostly used. For above scheme, If $\{\lambda_k\}$ remains bounded, then the penalty parameter sequence $\{c_k\}$ will remains bounded.

Another way is to use a different penalty parameter for each constraint $h_i(x) = 0$, and to increase the penalty parameter by a certain factor which correspond to those equations for which the constraint violation, measured by $|h_i[x(\lambda_k, c_k)]|$ is not decreased by a factor over the minimization.

Thus, from above discussion it can be noted that the method strongly depends on the initial choice of the multiplier $\lambda_0$ and choice of $\lambda_0$ close to $\lambda^*$ will give best convergence result and reduces computational requirement of the method.

## REFERENCES

[1]. Bertsekas, D.P. (1973) Convergence rate of penalty and multiplier methods. Proc.1973 IEEE Confer. Decision Control, San Diego, Calif., pp. 260-264.

[2]. Bertsekas, D.P. (1975a). Multiplier methods for convex programming. IEEE Trans. Automat. Control **20**, 385-388.

[3]. Bertsekas, D.P. (1975b). Necessary and sufficient conditions for a penalty method to be exact. M.P.**9**.87-99.

[4]. Bertsekas, D.P. (1976a). On penalty and multiplier methods for constrained optimization. S.J.C.O.**14**, 216-235.

[5]. Bertsekas, D.P. (1976b). Multiplier methods: A survey. Automatica-J.IFAC **12**, 133-145.

[6]. Bertsekas, D.P. (1977). Approximation procedures based on method of multipliers. J.O.T.A. **23,** 487-510.

[7]. Bertsekas, D.P. (1978). On the convergence properties of second order methods of multipliers. J.O.T.A. **25,** 443-449.

[8]. Bertsekas, D.P. (1979a). A convergence analysis of the method of multipliers for nonconvex constrained optimization. Presented at Proc. Workshop Augumented Lagrangians, IIASA. Vienna.

[9]. Biggs, M.C. (1978). On the convergence of some constrained minimization algorithms based on recursive quadratic programming. J. Inst. Math. Appl. **21,** 67-81.

[10]. Boggs, P.T., and Tolle, J.W. (1980). Augumented Lagrangian which are quadratic in the multiplier. J.O.T.A. **31,** 17-26.